DOCUMENT RESUME

ED 359 937 IR 016 157

AUTHOR Stanton, Michael J.
TITLE Digital Audio Sampling for Film and Video.
PUB DATE [Mar 93]
NOTE 12p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Acoustics; *Audio Equipment; Educational Technology; *Films; *Interactive Video; *Technological Advancement
IDENTIFIERS Analog Transmission Systems; Digital Data Signals

ABSTRACT

Digital audio sampling is explained, and some of its implications in digital sound applications are discussed. Digital sound equipment is rapidly replacing analog recording devices as the state-of-the-art in audio technology. The philosophy of digital recording involves doing away with the continuously variable analog waveforms and turning the patterns into numbers. A digital recording device rapidly samples the incoming sounds, quantifying the signal into a series of numerical values (binary codes). Although digital sound eliminates many of the traditional analog problems, digital signal processing presents key problems in sampling rates and synchronization. Careful control is necessary to check signals through each step in the audio chain. The advantages of digital audio processing include increased signal-to-noise ratio, no flutter, transparent generation of copies, and sound manipulation. These benefits come with increased quality control requirements. One chart illustrates sampling formats. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED359937

Digital Audio Sampling for Film and Video

Michael J. Stanton, Ph.D.

California State University, Northridge

Radio-Television-Film Department

120161577

"PERMISSION TO REPRODUCE THIS MATERIAL IN **OTHER THAN PAPER COPY** HAS BEEN GRANTED BY

Michael J. Stanton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

BEST COPY AVAILABLE

Digital Audio Sampling for Film and Video

Introduction

Current film and video production and post-production employ digital audio. Digital sound equipment is rapidly replacing analog recording devices as the "state-of-the-art" in audio technology. Therefore, personnel involved in the business of sound reproduction and manipulation, as well as those in the academic community, need to know elements of digital audio sampling. The purpose of this paper is to explain digital audio sampling and some of its implications in digital sound applications.

Although the theoretical beginnings of digital audio processing can be traced to the 1930's in the work of researchers such as Nyquist and Shannon, common use of digital signal processing (DSP) began during the 1980's. The development of the compact disc (CD) promoted digital audio as the popular pinnacle of sound technology during this period. Matching the consumer popularity of CD's, the film and video industries also gradually embraced digital audio technology. Unfortunately, the idyllic plateau achieved by digital sound is not a total Utopia in all sound applications. Problems exist when implementing digital signal processing in film and video production and studio operations.

Analog Versus Digital

Digital sound challenges the traditional analog recording processes that have been used since the late nineteenth century. The work of Edison, Pathe' and Berliner formed the foundations of analog sound recording, but digital sound recording techniques challenged basic concepts about the nature of audio. Analog recording technology is based on the capture of a continually varying voltage from a transducer. This continuously variable voltage changes in time and contains various frequency

complexes. The ability to capture all of the sounds in the original signal, the frequency response, proved adequate in the analog environment. Typical analog frequency responses while recording on a reel-to-reel machine at 15 ips (inches-per-second) ranged between 13-19 kHz (kilohertz) within 3 db (decibels) and a signal-to-noise ratio without noise reduction of about 60-65 dbs. These levels were considered commendable by the recording industry, but these capture rates of sound are far from what is currently available through digital recording techniques.

A major disadvantage of analog recording is that the patterns, or sound waveforms, first captured in an analog recording cannot be duplicated exactly in a re-recording of the same material. The mismatch between the original pattern and the pattern of the second generation copy results in so-called "artifacts" in the copy which account for audible noise and hiss. For example, when copying or editing the original material, approximately 3 db of contaminants are added each succeeding time the material passes through a tape recorder. Besides the additive noise, analog recording has other drawbacks. Analog recordings suffer from noise, distortion, and less than ideal headroom at high frequencies.

Digital recording technology bypasses most of analog's shortcomings. The philosophy of digital recording involves doing away with the continuously variable analog waveforms and turns the patterns into numbers. A digital recording device rapidly samples the incoming sounds, quantifying the signal into a series of numerical values (binary codes of "1's" and "0's"). This series of digits can be recovered with much better signal-to-noise ratio and much lower distortion than in analog recordings. Typical digital frequency responses range between 20-22 kHz within .5 db and exhibit a signal-to-noise ratio of 94-96 db at all levels. Distortion is very low, and wow and flutter are reduced. In their numerical form digital signals can be duplicated over and over again with incredible accuracy while remaining clear and clean.

Computer Workstations

Although digital signals can be stored on audio or video tape, digital audio developed simultaneously with computer technologies. Digital audio is in fact an outgrowth of computers and culminates in the computer workstation, a device capable of controlling a variety of studio audio equipment. Computer workstations may control audio recorders or may act as recorders themselves as sound is captured on a computer hard disc. Once captured within the digital domain, the signals can be copied and edited. Signals can be modified in many ways with great precision. Consider equalization (EQ) for example. A particular frequency band, e.g. a 50 Hz "main" hum indigenous to recording material in Europe, can be precisely removed in the digital domain. Signals can also be mixed emulating a multi-track, full-function studio mixer within a single digital workstation. In fact, signals can be manipulated in nearly infinite ways. One example of such manipulation with a computer workstation is time compression or expansion. Once the numerical values of the signals are converted into bits with a computer, the material can be shortened or lengthened without shifts in pitch. A thirty-two second commercial can be made to fit inconspicuously into a normal thirty second commercial window.

Based on the positive advantages in duplication, editing, manipulation and mixing, the digital environment appears to be a flawless world. The utility of digital sound makes it a godsend for film and video. Yet, although digital sound eliminates many of the traditional analog audio problems, digital signal processing presents some new problems of its own. The key hurdles in digital recording revolve around the issue of digital recording sampling rates and synchronization.

Digital Sampling Rates

In 1928, Harry Nyquist, a Bell Laboratories engineer, determined that for the highest frequency in a signal to be successfully encoded digitally, the signal had to be sampled at a rate at least twice its highest frequency. Therefore, a sampling frequency of at least 40 kHz will be required to capture frequencies as high as 20 kHz (footnote #1). This concept retains the name of its discoverer and is referred to as the Nyquist sampling theorem.

Sampling can be defined as rapid measurements of the values of audio signals and the conversion to binary digital data. This digital data is then recorded to some storage medium such as tape or computer memory. No single sampling practice has been established. In fact, the recording industries employ several different sampling rates. The sampling rate of compact discs is 44.1 kHz (44,100 Hz), meaning that the highest sampled audio can be approximately 22 kHz. Before the Phillips Corporation of the Netherlands realized the possibilities of storing audio signals optically in 1974 (footnote #2), videotape was considered for storage due to its large bandwidth capabilities. For example, a 44.1 kHz signal using a "word length" of 16-bits in an analog to digital (A to D) conversion yields 705,600 bits per second (44,100 x 16), a great deal of information to be stored. As indicated by Ken Pohlmann in his book Principles of Digital Audio, "the large amounts of data contained in even short musical selections place great demands on even the most sophisticated data storage and processing systems." (footnote #3, p. 229)

The normal NTSC (National Television System Committee) video signal used in the United States and Japan became one basis for digital audio sampling. Fitting audio data into video frames is primarily determined by the video scan rate. For a standard NTSC video signal a video frame occupies 1/30 of a second. Therefore, the

Quantized audio data is placed on each horizontal scanning line. Within normal NTSC video with its 525 scanning lines, thirty-five of the 525 scanning lines are used for vertical blanking leaving 490 lines for video information per video frame. Due to interlace or interleave scanning, each video frame is divided into an odd and even field, or half of the video information. This means there are 245 lines per field. With three audio samples per line provides for sufficient frequencies as per the Nyquist theorem, that is a video field rate of 60 times 245 times 3 equals 44.1 k (60 x 245 x 3 = 44,100). It also happens that 44.1 kHz is a convenient sampling rate for the European video standard PAL (Phase Alternation Line) which scans 625 lines each frame. With PAL a field rate of 50 is multiplied by 294, the available number of lines per field, times 3 to also arrive at 44.1 kHz (50 x 294 x 3 = 44,100). Although 44.1 kHz became the standard for compact discs, this sampling rate also served as a quasi-standard for video to successfully capture digital audio on videotape.

Although it would appear that a single standard existed for video and CD technology, the first digital audio processors for video did not use the CD sampling rate of 44.1 kHz. The first digital audio processors introduced in 1977 were designed to have a sampling rate of 40.056 kHz. This rate was derived from the exact NTSC "color time" of 59.94 fields per second (59.94 x 245 x 3 = 44,055.9), and 49.95 fields per second for PAL (49.95 x 294 x 3 = 44,055.9). This sampling rate of 44.056 kHz was created simply because it worked directly with the videotape recorders already manufactured by the companies involved. Again, videotape had a large enough bandwidth to store the large amounts of data generated by digital sampling. So, at the start of the 1980's, two sampling rates existed with both standards in general use within film and video production. By 1982, another digital sampling rate was introduced to the professional recording industry. This rate was the Audio Engineering Society (AES) 48 kHz sampling frequency, a digital audio standard for open reel audio

tape machines used primarily for sound mastering. The idea behind this rate was to provide an even greater spectrum of sound than what would be eventually delivered to a customer. In this way recordings originally produced at 48 kHz would eventually be stepped down to either 44.056 kHz or 44.1 kHz. Currently the 48 kHz sampling rate is used by several types of recorders. D-2 and DAT machines can record at 48 kHz.

The relationship between the sampling frequency and NTSC line frequencies, and any other sampling frequencies is largely immaterial. Recorders and converters exist so that one sampled frequency can be converted to any other frequency. Digital audio interfaces exist to link the various equipment from the same manufacturer as well as between hardware from different manufacturers. Through digital interfacing, the signals can be preserved and transferred through the audio chain from initial source to ultimate playback.

Digital Recording in Real Situations

Although digital audio conversions can be conducted with no appreciable changes in digital audio signals, problems arise in real world video and film production and post-production. These situations stem from poor record keeping and quality control resulting in a lack of synchronization between picture and sound. The primary concern is the so-called time-alignment of sample "clocks" within digital audio signal sources. The offsets between sound and any corresponding pictures are akin to the difficulties that may happen in transferring between the 24 frames per second of film and the 30 frames per second of video. In other words, the fact that different digital audio sampling rates exist may cause a lack of synchronization.

At a recent SMPTE conference on digital audio, Jeff Evans of Audio Intervisual Design recounted a true story exemplifying what can happen in film or video projects. According to Mr. Evans,

> It was a project that was a film that was shot in England. They had three recorders and a digital multi-track with no reference code. They shot the picture. They did some more work on the audio on multi-track. It was converted to PAL at this point for workprints. It came to the U.S. and was converted to NTSC. They tried more multi-track work, and they wanted the U.S. release version on film. So, we went from 50 to PAL to NTSC to 24 with digital audio moving all the way through it. Guess what? It didn't work. (footnote #4)

In this case the synchronization between picture and sound was lost. Digital interconnections involved equipment, either local or remote, without timing under control of the studio or studio center. The mismatch was due to inattention to conversion details such as maintaining reference points and failure to adjust for sample per frame rates.

According to the Audio Engineering Society (AES), equipment should be synchronized by one of two methods. The preferred method for normal studio practice is the use of the Digital Audio Reference Signal which ensures that all input-output equipment sample clocks are locked to that reference. The second method is through the use of the embedded sample rate clock within a digital audio input signal, that enables input-output sample rate clocks to be locked. This method is less desirable because it may increase the timing error between items of equipment in a cascaded implementation. (footnote #5, p. 4)

Each sampling rate is based on a different sample per frame rate and offsets may occur from one rate to another. In the case of a combined video and audio environment, the source of the Digital Audio Reference Signal must be locked to the video source so that the mathematical relationships given below may be obtained

precisely. (footnote # 6, p. 5)

[Insert "Common Digital Audio Sampling Formats" Chart Here]

The above chart was created by Audio Intervisual Design to catalog common digital audio formats. Based on the number of digital audio samples measured per frame of media. For example, an NTSC signal may be sampled at 48 kHz so that a 48 kHz digital track can be synchronized with it. For every video frame, there will be 1601.6 audio "frames," i.e. 48000/29.97 or the sampling rate divided by the frame rate. Assuming a video switcher generates switching pulses at every video frame boundary, then one video frame later exactly 1601.6 audio frames will have been measured. Switching both of these signals at this point means a switch .6 of the way through an audio frame. An offset has occurred and a series of offsets will continue through further switching or editing. In this way routing and distributing digital audio signals can become quite unpredictable. Although analog devices may be interconnected with ease, transferring digital audio signals from one format to another is more complex. Sampling rates, reference points, "word" lengths, and synchronization must be precisely monitored to grant errorless interfacing.

Numerous data formats exist to permit the interconnection of digital audio devices, whether the equipment comes from the same manufacturer or several different manufacturers. Several digital audio interface standards exist and among them are SDIF-2 (Sony Digital Interface Format), the PD (ProDigi) interface, the AES/EBU (Audio Engineering Society/European Broadcasting Union) interface, the International Electrotechnical Commission (IEC) interface, and the multichannel audio digital interface (MADI). Although these different and frequently incompatible standards are available in commercial equipment, only through digital interfacing can

fidelity and synchronization of digital signals be preserved through the audio production chain.

Conclusions

Any project involving digital audio processing must always keep the final output format in mind. Conversions can be conducted from format to format, but digital audio requires new levels of quality control and careful monitoring of a project through all digital transfer stages. References must also be kept intact. Notes need to be studiously kept on all material through each stage of production and post-production. Simple tasks such as labeling tapes takes on a heightened importance with digital audio.

To summarize, the advantages of digital audio signal processing are indeed obvious: increased signal-to-noise ratios, no flutter, "transparent" generations of copies, and sound manipulation among others. Yet, these benefits come with associated costs including increased quality control diligence in checking signals through each step in the audio chain.

COMMON DIGITAL AUDIO SAMPLING FORMATS

	Sample Rate (Hz.)	Frame Rate	Samples per Frame
Music:			
CD	44100	30	1470 (44100/30)
CD (Video Speed)	44056	29.97	1470 (44056/29.97)
	48000	30	1600 (48000/30)
Film:			
ORC	44100	24	1837.5 (44100/24)
(Optical Radiation Corporation)			
Dolby	48000	24	2000 (48000/24)
Video:			
NTSC	44056	29.97	1470 (44056/29.97)
	48000	29.97	1601.6 (48000/29.97)
Video:			
PAL	48000	25	1920 (48000/25)
Video:			
HDTV	48000	30	1600 (48000/30) (footnote #7)